

SALES ANALYSIS ON BIG BASKET PRODUCTS DATA SET

DOMAIN: ECOMMERCE WEBSITES

TOOLS USED: PYTHON FOR DATA CLEANING AND PRE-PROCESSING
APACHE PIG FOR DATA ANALYSIS

DATASOURCE: KAGGLE

<https://www.kaggle.com/hetulmehta/bigbasket-products-dataset>

SCHEMA:

PRODUCT, CATEGORY, SUB-CATEGORY, BRAND, TYPE, RATING,
SALE_PRICE, MARKET_PRICE, PROFIT, PROFIT%

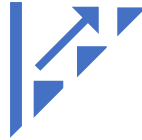
SIZE: 22.45 MB



DATA CLEANING IN PYTHON

- UNECESSARY COLUMNS LIKE URLS, CODES WERE REMOVED
- NULL VALUES AND DUPLICATES WERE REMOVED
- MISSING VALUES IN RATING COLUMNS WE REPLACED WITH ITS MEDIAN.

DETAILED EXPLANATION IN JUPYTER NOTEBOOK



ANALYSIS PART USING PIG

1) LOAD, DESCRIBE AND DUMP OPERATIONS

```
kevin@kevin-VirtualBox: ~/Desktop
grunt> project = LOAD '/home/kevin/Desktop/new.csv' using PigStorage(',') as (product : chararray , category: chararray , sub_category: chararray ,brand:chararray,sale_price:float,market_price:float,type:chararray,rating:float,Profit:float,Profitpercentage:float);
2021-10-03 10:03:03,141 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
grunt> describe project;
project: {product: chararray,category: chararray,sub_category: chararray,brand: chararray,sale_price: float,market_price: float,type: chararray,rating: float,Profit: float,Profitpercentage: float}
grunt> dump project;
```

DUMPED DATA

```
Terminal
Oct 3 10:04
kevin@kevin-VirtualBox: ~/Desktop

(Almond Oil,Beauty & Hygiene,Skin Care,Aroma Magic ,300.0,400.0,Aromatherapy,2.0,100.0,33.3)
(3-In-1 Face Wash,Beauty & Hygiene,Skin Care,Organic Harvest,182.75,215.0,Face Care,4.0,32.25,17.6)
(Radiant Pearl Cellular Lightening 1 Facial Kit,Beauty & Hygiene,Skin Care,Lotus Herbals,171.5,245.0,Face Care,4.5,73.5,42.9)
(Tea Tree Skin Clearing Scrub,Beauty & Hygiene,Skin Care,BIOTIQUE,142.5,190.0,Face Care,4.1,47.5,33.3)
(Gold Creme Bleach,Beauty & Hygiene,Skin Care,Natures Essence,70.0,70.0,Face Care,3.3,0.0,0.0)
(Skin Fruits Softening Glow Apple Face Wash,Beauty & Hygiene,Skin Care,Joy,97.75,115.0,Face Care,5.0,17.25,17.6)
(Sunscreen Lotion - Carrot,Beauty & Hygiene,Skin Care,Aroma Magic ,148.75,175.0,Face Care,3.7,26.25,17.6)
(Real Nature Tea Tree Face Mask,Beauty & Hygiene,Skin Care,The Face Shop,100.0,100.0,Face Care,4.2,0.0,0.0)
(Dreamcatcher Essential Oil,Beauty & Hygiene,Skin Care,Soulflower,360.0,450.0,Aromatherapy,4.8,90.0,25.0)
(Charcoal Face Scrub,Beauty & Hygiene,Skin Care,INATUR ,315.0,450.0,Face Care,5.0,135.0,42.9)
(Charcoal Peel-Off Face Mask,Beauty & Hygiene,Skin Care,INATUR ,360.0,450.0,Face Care,3.6,90.0,25.0)
(Natural Exfoliating Face Wash For Men & Women,Beauty & Hygiene,Skin Care,Sirona,299.0,375.0,Face Care,4.0,76.0,25.4)
(Cleanser-Moisturiser Combo,Beauty & Hygiene,Skin Care,Kaya Clinic,1190.0,1190.0,Face Care,4.1,0.0,0.0)
(Oxy-Infusion Face Serum,Beauty & Hygiene,Skin Care,Kaya Youth,349.3,499.0,Face Care,1.0,149.7,42.9)
(Activated Charcoal Face Wash - Skin Purifying + Deep Detox,Beauty & Hygiene,Skin Care,Joy,136.5,195.0,Face Care,4.4,58.5,42.9)
(Skin Lightening Shiner Bleach With Diamond Dust,Beauty & Hygiene,Skin Care,Olivvia,17.0,20.0,Face Care,4.3,3.0,17.6)
(Aloe Vera Gel,Beauty & Hygiene,Skin Care,ST. D' VENCE,399.0,399.0,Face Care,3.8,0.0,0.0)
(Essential Oil - Jasmine,Beauty & Hygiene,Skin Care,Aroma Magic ,375.0,500.0,Aromatherapy,5.0,125.0,33.3)
(Rejuvenating Night Cream - Cucumber Lime & Aloe Vera,Beauty & Hygiene,Skin Care,Sri Sri Ayurveda ,70.0,70.0,Face Care,4.0,0.0,0.0)
("Sun Shield - Bio Pink Lentils Matte Look Sunblock, SPF 40",Beauty & Hygiene,Skin Care,,296.25,395.0,,3.0,98.75)
(Vitamin C Brightening Day Cream With SPF 30 UVA/UVB PA+++,Beauty & Hygiene,Skin Care,StBotanica,669.0,669.0,Face Care,5.0,0.0,0.0)
(Hydro Replenish Gentle Face Wash,Beauty & Hygiene,Skin Care,Kaya Youth,139.3,199.0,Face Care,3.7,59.7,42.9)
(Peppermint & Black Seed Lip Polish,Beauty & Hygiene,Skin Care,INATUR ,225.0,300.0,Lip Care,3.0,75.0,33.3)
(Divine Nutritive Cream SPF 20,Beauty & Hygiene,Skin Care,Lotus Organics+,616.25,725.0,Face Care,4.1,108.75,17.6)
(Pristine Purifying Face Wash,Beauty & Hygiene,Skin Care,Lotus Organics+,293.25,345.0,Face Care,4.1,51.75,17.6)
(CoCo Face Mask With Coffee & Cocoa For Glowing Skin,Beauty & Hygiene,Skin Care,Mamaearth,499.0,499.0,Face Care,4.0,0.0,0.0)
(Natural Clay Face Wash,Beauty & Hygiene,Skin Care,The Moms Co,393.0,393.0,Face Care,4.5,0.0,0.0)
(Protecting Gel Face Wash,Beauty & Hygiene,Skin Care,Natures Essence,65.0,65.0,Face Care,3.8,0.0,0.0)
(Moisturizing Cream,Beauty & Hygiene,Skin Care,Sebamed,445.0,559.0,Face Care,4.3,114.0,25.6)
(Earth C7 Argan Oil,Beauty & Hygiene,Skin Care,ColorBar,760.0,800.0,Aromatherapy,5.0,40.0,5.3)
(Tea Tree Anti Imperfection Daily Solution,Beauty & Hygiene,Skin Care,BIOTIQUE,420.0,560.0,Face Care,4.1,140.0,33.3)
(Lip Butter - Rose,Beauty & Hygiene,Skin Care,Organic Harvest,169.15,199.0,Lip Care,1.5,29.85,17.6)
(Intimate Lightening Serum,Beauty & Hygiene,Skin Care,Namyaa,750.0,750.0,Intimate Wash & Care,4.0,0.0,0.0)
(Wine Bleach,Beauty & Hygiene,Skin Care,Astaberry,71.25,75.0,Face Care,3.4,3.75,5.3)
(Pure Aloe Anti Pollution Face Wash,Beauty & Hygiene,Skin Care,Joy,82.5,110.0,Face Care,4.4,27.5,33.3)
(Mulberry & Rose Facewash,Beauty & Hygiene,Skin Care,Vlcc,171.0,190.0,Face Care,4.0,19.0,11.1)
(Cleanser - Mint,Beauty & Hygiene,Skin Care,Aroma Magic ,385.0,550.0,Face Care,3.3,165.0,42.9)
(Scrub - Coffee Bean,Beauty & Hygiene,Skin Care,Aroma Magic ,281.25,375.0,Face Care,4.8,93.75,33.3)
(Fair+Glow Sun Screen Lotion SPF 20,Beauty & Hygiene,Skin Care,Vlcc,178.75,275.0,Face Care,4.7,96.25,53.8)
(Precious Brightening Face Wash,Beauty & Hygiene,Skin Care,Lotus Organics+,408.0,480.0,Face Care,3.0,72.0,17.6)
(Skin Healing Sheet Mask (Pack of 6),Beauty & Hygiene,Skin Care,Mirabelle,392.73,714.0,Face Care,4.1,321.27,81.8)
(Youthrx Anti Ageing Firming Face Masque,Beauty & Hygiene,Skin Care,Lotus Herbals,236.0,295.0,Face Care,3.3,59.0,25.0)
(Natural Baby Cream,Beauty & Hygiene,Skin Care,The Moms Co,393.0,393.0,Face Care,4.3,0.0,0.0)
(Apple Cider Vinegar Imperfection Daily Solution,Beauty & Hygiene,Skin Care,BIOTIQUE,420.0,560.0,Face Care,3.0,140.0,33.3)
("Gold Glow Face Wash For Radiant Skin With Gold, Chandan, Turmeric",Beauty & Hygiene,,12.75,15.0,,4.0)
(Multani Mati,Beauty & Hygiene,Skin Care,Satinance,58.0,58.0,Face Care,3.6,0.0,0.0)
(Ultra Light Aquagel Sunscreen SPF 25,Beauty & Hygiene,Skin Care,Kaya Clinic,406.0,580.0,Face Care,3.5,174.0,42.9)
(Lighten & Smooth Under Eye Gel,Beauty & Hygiene,Skin Care,Kaya Clinic,343.0,490.0,Eye Care,2.8,147.0,42.9)
(Radiant Pearl Cellular Lightening 1 Facial Kit,Beauty & Hygiene,Skin Care,Lotus Herbals,171.5,245.0,Face Care,4.5,73.5,42.9)
```

QUERIES:

1) MAXIMUM RATED PRODUCTS: THAT IS PRODUCTS WITH RATING 5

```
(Premium Adult Diaper Disposal Bags,Beauty & Hygiene,Feminine Hygiene,Sirona,297.5,350.0,Intimate Wash & Care,3.0,32.5,17.6)
grunt> temp1 = FILTER project by rating==5.0;
2021-10-03 10:18:01,491 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 2 time(s).
grunt> dump temp1;
```

OUTPUT:

```
es Terminal Oct 3 10:20
kevin@kevin-VirtualBox: ~/Desktop

(Baby Happy Tissues,Baby Care,Diapers & Wipes,Bella,144.0,150.0,Baby Wipes,5.0,0.0,4.2)
(Easy Adult Diaper - Large,Baby Care,Diapers & Wipes,Friends,451.5,525.0,Diapers,5.0,73.5,16.3)
(Dry Pants Diapers - X-Large,Baby Care,Diapers & Wipes,Huggies,319.33,349.0,Diapers,5.0,29.67,9.3)
(Baby Happy Wet Wipes - Milk & Honey,Baby Care,Diapers & Wipes,Bella,189.0,189.0,Baby Wipes,5.0,0.0,0.0)
(Premium Soft Diaper Pants - Extra Small,Baby Care,Diapers & Wipes,Huggies,1091.09,1199.0,Diapers,5.0,107.91,9.9)
(Wonder Pants Large - 99 Diapers,Baby Care,Diapers & Wipes,Huggies,1463.09,1599.0,Diapers,5.0,135.91,9.3)
(Total Care Baby Pants Diapers - S,Baby Care,Diapers & Wipes,Himalaya,1000.0,1000.0,Diapers,5.0,0.0,0.0)
(Baby Happy Two-Layer Universal Tissues,Baby Care,Diapers & Wipes,Bella,417.0,417.0,Baby Wipes,5.0,0.0,0.0)
(Baby Happy Large - 48 Diapers,Baby Care,Diapers & Wipes,Bella,739.0,739.0,Diapers,5.0,0.0,0.0)
(Baby Wipes - Combo Pack,Baby Care,Diapers & Wipes,Morison Baby Dreams,258.0,258.0,Baby Wipes,5.0,0.0,0.0)
(Reusable Diaper Insert Pads,Baby Care,Diapers & Wipes,Paw Paw,339.0,339.0,Diapers,5.0,0.0,0.0)
(Milo,Gourmet & World Food,Drinks & Beverages,Nestle ,550.0,550.0,Health Drinks,5.0,0.0,0.0)
(Nesquik - Strawberry,Gourmet & World Food,Drinks & Beverages,Nestle ,550.0,550.0,Health Drinks,5.0,0.0,0.0)
(Nesquik - Chocolate,Gourmet & World Food,Drinks & Beverages,Nestle ,550.0,550.0,Health Drinks,5.0,0.0,0.0)
(Liver Cleanse Tea,Gourmet & World Food,Drinks & Beverages,Nanhya,589.0,589.0,Health Drinks,5.0,0.0,0.0)
(PCOS Tea,Gourmet & World Food,Drinks & Beverages,Nanhya,589.0,589.0,Health Drinks,5.0,0.0,0.0)
(Sweet Flavour Sattuz,Gourmet & World Food,Drinks & Beverages,Sattuz,120.0,120.0,Health Drinks,5.0,0.0,0.0)
(Spiced Turmeric Latte Mix,Gourmet & World Food,Drinks & Beverages,Rooted Peepul,500.5,599.0,Health Drinks,5.0,29.5,5.3)
(Diabetes Care Tea,Gourmet & World Food,Drinks & Beverages,Nanhya,589.0,589.0,Health Drinks,5.0,0.0,0.0)
(Chocolate Flavour Sattuz,Gourmet & World Food,Drinks & Beverages,Sattuz,150.0,150.0,Health Drinks,5.0,0.0,0.0)
(Chocolate Flavour Sattuz,Gourmet & World Food,Drinks & Beverages,Sattuz,360.0,360.0,Health Drinks,5.0,0.0,0.0)
(Energy Drink Beverages,Energy & Soft Drinks,Red Bull,3480.0,3480.0,Sports & Energy Drinks,5.0,0.0,0.0)
(Non-Alcoholic Beer Bottle - Malt Flavour Beverages,Energy & Soft Drinks,Barbicam ,99.0,99.0,Non Alcoholic Drinks,5.0,0.0,0.0)
(FrutORS - Apple,Beverages,Energy & Soft Drinks,Real,35.0,35.0,Sports & Energy Drinks,5.0,0.0,0.0)
(Easy Grip 360-Degree Trainer Sipper Cup - Blue,Baby Care,Feeding & Nursing,Mee Mee,399.0,399.0,Baby Dishes & Utensils,5.0,0.0,0.0)
(Standard Neck Nursing Bottle KPP With S-Type Nipple - Light Blue,Baby Care,Feeding & Nursing,Pigeon Baby,255.0,255.0,Sippers & Bottles,5.0,0.0,0.0)
(Ultra Absorbent Disposable Period Panties for Sanitary Protection - S-M,Beauty & Hygiene,Feminine Hygiene,Eve,Eve,50.0,100.0,Sanitary Napkins,5.0,50.0,100.0)
(Regimen Kit - 1 Razor + 2 Refills + 1 Satin Care Sensitive Skin Gel,Beauty & Hygiene,Feminine Hygiene,Gillette Venus,799.2,999.0,Hair Removal,5.0,199.8,25.0)
(Hair Removal Waxing Kit - Normal Skin,Beauty & Hygiene,Feminine Hygiene,Veet,179.1,199.0,Hair Removal,5.0,19.9,11.1)
(Venus Snap Hair Remover For Smooth Skin - Women,Beauty & Hygiene,Feminine Hygiene,Gillette,799.0,799.0,Hair Removal,5.0,0.0,0.0)
(Panty Liners - Intima Plus Extra Long,Beauty & Hygiene,Feminine Hygiene,Bella,249.0,249.0,Panty Liners,5.0,0.0,0.0)
(100% Pure US Cotton Ultra Thin Sanitary Pads - Travel Pack,Beauty & Hygiene,Feminine Hygiene,Plush,25.0,25.0,Sanitary Napkins,5.0,0.0,0.0)
(Premium Digital Tampon - Heavy Flow,Beauty & Hygiene,Feminine Hygiene,Sirona,146.2,215.0,Tampons & Menstrual Cups,5.0,68.8,47.1)
(Perfecta Ultra Silky Drai Sanitary Napkins,Beauty & Hygiene,Feminine Hygiene,Bella,150.0,150.0,Sanitary Napkins,5.0,0.0,0.0)
(Premium Applicator Tampons - Regular Flow,Beauty & Hygiene,Feminine Hygiene,Sirona,403.2,630.0,Tampons & Menstrual Cups,5.0,226.8,56.2)
(Refreshing Intimate Wash - Natural,Beauty & Hygiene,Feminine Hygiene,Sirona,169.15,199.0,Intimate Wash & Care,5.0,29.85,17.6)
(Panty Liners - Intima Plus Normal,Beauty & Hygiene,Feminine Hygiene,Bella,239.0,239.0,Panty Liners,5.0,0.0,0.0)
(Premium Digital Tampon - Heavy Flow,Beauty & Hygiene,Feminine Hygiene,Sirona,227.01,329.0,Tampons & Menstrual Cups,5.0,101.99,44.9)
(Portable Female Urination Device For Women - Disposable,Beauty & Hygiene,Feminine Hygiene,Peebuddy,276.25,325.0,Intimate Wash & Care,5.0,48.75,17.6)
(Clean & Pure Intimate Hygiene Wash With Tea Tree Oil & Sea Buckthorn,Beauty & Hygiene,Feminine Hygiene,Dr. Morepen,161.0,175.0,Intimate Wash & Care,5.0,14.0,8.7)
(Premium Adult Diaper Disposal Bags,Beauty & Hygiene,Feminine Hygiene,Sirona,148.75,175.0,Intimate Wash & Care,5.0,26.25,17.6)
(After Waxing Oil - Lavender,Beauty & Hygiene,Feminine Hygiene,Jeva,337.5,450.0,Hair Removal,5.0,112.5,33.3)
(Liposoluble Wax - Banana,Beauty & Hygiene,Feminine Hygiene,Jeva,712.5,950.0,Hair Removal,5.0,237.5,33.3)
(Tampons - Regular,Beauty & Hygiene,Feminine Hygiene,Floah,115.0,115.0,Tampons & Menstrual Cups,5.0,0.0,0.0)
(Premium Adult Diaper Disposal Bags,Beauty & Hygiene,Feminine Hygiene,Sirona,297.5,350.0,Intimate Wash & Care,5.0,52.5,17.6)
```

FROM THE OUTPUT WE CAN SEE THAT THE HIGHEST RATED PRODUCTS ARE : DRINKS(TEA) , SANITARY PRODUCTS AND PRODUCTS FOR BABIES LIKE TISSUES, DIAPERS ETC.

2) MINIMUM RATED PRODUCTS: THAT IS PRODUCTS WITH RATING 1

```
grunt> temp2 = FILTER project by rating==1.0;
2021-10-03 10:20:55,013 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPL
ICIT_CAST_TO_DOUBLE 3 time(s).
grunt> dump temp2
```

OUTPUT:


```
(Acnoklin-Rs Gel Face Wash,Beauty & Hygiene,Skin Care,Vegetal,220.0,220.0,Face Care,1.0,0.0,0.0)
(Massage Oil - Apricot Kernel Oil,Beauty & Hygiene,Skin Care,Aloe Veda,472.5,525.0,Aromatherapy,1.0,52.5,11.1)
(Sunscreen SPF 30,Beauty & Hygiene,Skin Care,Organic Harvest,420.75,495.0,Face Care,1.0,74.25,17.6)
(Glow Oxy Bleach Crene,Beauty & Hygiene,Skin Care,Globus,192.0,240.0,Face Care,1.0,48.0,25.0)
(Retinol Advanced Anti Aging Face Mist,Beauty & Hygiene,Skin Care,StBotanica,699.0,699.0,Face Care,1.0,0.0,0.0)
(Coconut Oil,Beauty & Hygiene,Skin Care,St. D'VENCE,254.0,299.0,Aromatherapy,1.0,45.0,17.7)
(Urban Balance Daily Scrub Cleanser,Beauty & Hygiene,Skin Care,Faces,411.75,549.0,Face Care,1.0,137.25,33.3)
(Face Cream - Fairness,Beauty & Hygiene,Skin Care,Vegetal,260.0,260.0,Face Care,1.0,0.0,0.0)
(Body Moisturising Lotion,Beauty & Hygiene,Skin Care,Vegetal,340.0,340.0,Body Care,1.0,0.0,0.0)
(Body Massage Oil - Jasmine & Nargis,Beauty & Hygiene,Skin Care,Oriental Botanics,899.0,899.0,Aromatherapy,1.0,0.0,0.0)
(Gardenia Essential Oil,Beauty & Hygiene,Skin Care,Organic Harvest,311.6,410.0,Aromatherapy,1.0,98.4,31.6)
(Facial Kit - Pearl,Beauty & Hygiene,Skin Care,Arona Magic ,690.0,1150.0,Face Care,1.0,460.0,66.7)
(Gold Kit,Beauty & Hygiene,Skin Care,Natures Essence,1800.0,1800.0,Face Care,1.0,0.0,0.0)
(Diamond Kit,Beauty & Hygiene,Skin Care,Natures Essence,1800.0,1800.0,Face Care,1.0,0.0,0.0)
(Face and Body Mist - Frankincense,Beauty & Hygiene,Skin Care,ORGANIC INDIA,325.0,325.0,Body Care,1.0,0.0,0.0)
(Rich Hand V Special Care Hand Mask,Beauty & Hygiene,Skin Care,The Face Shop,212.5,250.0,Body Care,1.0,37.5,17.6)
(Merry Cherry Lip Balm - SPF 20,Beauty & Hygiene,Skin Care,BIOTIQUE,111.75,149.0,Lip Care,1.0,37.25,33.3)
(Caramel & Cedarwood Moisturizer,Beauty & Hygiene,Skin Care,RawNature,324.35,499.0,Face Care,1.0,174.65,53.8)
(Anti-Aging Night Cream,Beauty & Hygiene,Skin Care,INATUR ,337.5,450.0,Face Care,1.0,112.5,33.3)
(Fairy Oil,Beauty & Hygiene,Skin Care,Arona Magic ,300.0,400.0,Aromatherapy,1.0,100.0,33.3)
(Avocado Cold Pressed Oil,Beauty & Hygiene,Skin Care,Myasso,900.0,900.0,Aromatherapy,1.0,0.0,0.0)
(Rushikanti Multant Mitti With Rose,Beauty & Hygiene,Skin Care,Akshara Herbals,50.0,50.0,Face Care,1.0,0.0,0.0)
(Charcoal Detox Over Night Repair Face Cream,Beauty & Hygiene,Skin Care,INATUR ,350.0,500.0,Face Care,1.0,150.0,42.9)
(Ayurvedic Tea Tree & Activated Charcoal Face Pack,Beauty & Hygiene,Skin Care,Natural Vibes,499.0,499.0,Face Care,1.0,0.0,0.0)
(Aloe 90% Soothing Gel,Beauty & Hygiene,Skin Care,It's Skin,400.0,400.0,Face Care,1.0,0.0,0.0)
(Family Sunscreen Lotion SPF 25,Beauty & Hygiene,Skin Care,INATUR ,273.0,420.0,Face Care,1.0,147.0,53.8)
(Lip Butter - Chocolate,Beauty & Hygiene,Skin Care,Aloe Veda,225.0,225.0,Lip Care,1.0,0.0,0.0)
(Energy Essential Oil,Beauty & Hygiene,Skin Care,Soulflower,328.0,450.0,Aromatherapy,1.0,122.0,37.2)
(Herbal Gold Thermo Herb,Beauty & Hygiene,Skin Care,Khadi Natural,175.0,250.0,Face Care,1.0,75.0,42.9)
(Bananna Treat Chilll Masala,Snacks & Branded Foods,Snacks & Namkeen,FruitTreat,52.5,70.0,Namkeen & Savoury Snacks,1.0,17.5,33.3)
(Ragi Bites - Epic Masala,Snacks & Branded Foods,Snacks & Namkeen,Soulfull,20.0,20.0,Namkeen & Savoury Snacks,1.0,0.0,0.0)
(Disney Cinderella Princess PVC Embossed Pencil Bag - HHHTPB 72231-CIN,Cleaning & Household,Stationery,Hm International,100.0,199.0,Exam Pads & Pencil Box,1.0,39.0,24.4)
(Organics Indian Turmeric Tea,Beverages,Tea,10 Herbs,345.0,345.0,Leaf & Dust Tea,1.0,0.0,0.0)
(Darjeeling Green Tea - Ahimsa Long Leaf,Beverages,Tea,Teamonk,385.0,550.0,Exotic & Flavoured Tea,1.0,165.0,42.9)
(Sandwich Slices,Gourmet & World Food,Tinned & Processed Food,Entree,125.4,132.0,Tomatoes & Vegetables,1.0,6.6,5.3)
(Keto Pancake Mix,Gourmet & World Food,Cooking & Baking Needs,Ketofy,499.0,499.0,Flours & Pre-Mixes,1.0,0.0,0.0)
(Gluten-Free Vegetable Millet Khichdi Mix,Gourmet & World Food,Cooking & Baking Needs,Graininway,350.0,350.0,Flours & Pre-Mixes,1.0,0.0,0.0)
(Cream - Original,Gourmet & World Food,Dairy & Cheese,Nestle ,262.5,350.0,Milk & Soya Drinks,1.0,87.5,33.3)
(Wonder Pants Diapers - Small Size,Baby Care,Diapers & Wipes,Huggies,407.63,450.0,Diapers,1.0,42.37,10.4)
(Total Care Baby Pants Diapers - L,Baby Care,Diapers & Wipes,Himalaya,1065.0,1200.0,Diapers,1.0,135.0,12.7)
(Aloe Vera Juice - Purifies Blood & Boosts Immunity,Gourmet & World Food,Drinks & Beverages,Jiva Ayurveda,220.0,220.0,Health Drinks,1.0,0.0,0.0)
(Japanese Matcha Tea - 2 In 1 For Hot or Iced,Gourmet & World Food,Drinks & Beverages,Care,250.0,250.0,Health Drinks,1.0,0.0,0.0)
(Hibiscus & Red Melon Green Tea - 2 In 1 For Hot & Iced,Gourmet & World Food,Drinks & Beverages,Care,250.0,250.0,Health Drinks,1.0,0.0,0.0)
(Turtle Sipper - Green,Baby Care,Feeding & Nursing,Luvlap,245.0,245.0,Sippers & Bottles,1.0,0.0,0.0)
(Pro Super Soft Reusable FDA Approved Menstrual Cup With Medical Grade Silicon - Medium,Beauty & Hygiene,Feminine Hygiene,Sirona,719.1,799.0,Tampons & Menstrual Cups,1.0,79.9,11.1)
```

FROM THE OUTPUT WE CAN SEE THAT THE LEAST RATED PRODUCTS ARE : BEAUTY ,SKIN-CARE AND HYGEIENE TYPE PRODUCTS AND SOME FOOD ITEMS.

3) NUMBER OF PRODUCTS PER BRAND:

```
grunt> gbrand = GROUP project by brand;
2021-10-03 10:26:01,147 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> describe gbrand;
gbrand: {group: chararray,project: {(product: chararray,category: chararray,sub_category: chararray,brand: chararray,sale_price: float,market_price: float,discountpercentage: float)}}
grunt> cnt_prod = foreach gbrand generate group , COUNT(project);
2021-10-03 10:27:11,870 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
grunt> dump cnt_prod;
2021-10-03 10:27:20,573 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 3 time(s).
```

OUTPUT:

```

(Z,2)
(DP,11)
(Fa,3)
(HE,1)
(JK,1)
(KM,2)
(LG,1)
(Ob,1)
(Ok,1)
(Re,1)
(V2,2)
(iD,6)
(&Me,2)
(365,1)
(3B0,4)
(9t9,6)
(Aer,2)
(Axe,35)
(BGR,3)
(BSY,1)
(Cif,4)
(Cpp,1)
(DND,1)
(Dip,6)
(Ela,15)
(Eva,5)
(Exo,5)
(Fem,17)
(GPC,4)
(Grb,5)
(HIT,6)
(Hii,1)
(IK0,1)
(IN2,8)
(Iba,7)
(Ifb,1)
(Joy,39)
(KCL,3)
(Kvg,1)
(Lal,1)
(Lee,1)
(Lux,9)
(MOM,10)
(MTR,65)
(MVS,2)
(Mud,1)
(NSH,1)
(Neo,17)
(OCV,2)

```

FROM THE OUTPUT WE CAN SEE BRAND WISE COUNT OF PRODUCTS.

BRANDS like FRESHO , BBHOME ,COLGATE ETC.HAVE MORE NO. OF PRODUCTS.

4) TOTAL PROFIT MADE BY EACH BRAND

```
grunt> gbrand = GROUP project by brand;
2021-10-03 10:44:34,012 [main] WARN  org.apache
grunt> mpb = foreach gbrand {
>>   ordp = order project by Profit desc;
>>   maxpro = limit ordp 1;
>>   generate group , SUM(project.Profit);
>> };
2021-10-03 10:45:14,928 [main] WARN  org.apache
grunt> dump mpb;
```

OUTPUT:


```

(Al' Fez,0.0)
(Alkalen,0.0)
(All Out,94.0)
(Alvizia,476.0400085449219)
(Ambipur,0.0)
(Amrutha,7.400000095367432)
(Aramusk,445.0)
(Azafran,1528.9499998092651)
(BB Home,841.1000018119812)
(Bacto-V,0.0)
(Bagrrys,297.25)
(Bambino,0.0)
(Batiste,0.0)
(Battler,0.0)
(Beanies,140.0)
(Befikar,0.0)
(Bentley,2970.0)
(Bio-Oil,229.0)
(Bionova,0.0)
(Bolst's,0.0)
(Brimune,0.0)
(Britte ,0.0)
(Brother,38.75)
(Buffet ,0.0)
(Bvlgari,0.0)
(Cadbury,249.31999871134758)
(CamPure,0.0)
(Canopus,0.0)
(Carmesi,0.0)
(Cartini,57.99999952316284)
(Chaayos,0.0)
(Charmis,64.9000015258789)
(Cheetos,0.0)
(Cinthol,326.9600028991699)
(Clenz U,0.0)
(Closeup,90.0)
(Cocojal,0.0)
(Colgate,2918.3999996185303)
(Comfort,35.100000858306885)
(D-Alive,3.5)
(Daadis ,12.699999809265137)
(Daarzel,91.35000038146973)
(Delishh,4.799999952316284)
(Dexolac,0.0)
(Dheepam,0.0)
(Diamond,1.0)
(Dimples,0.0)
(Doritos,17.0)
(Et. Paso,0.0)

```

TOTAL PROFIT MADE BY EACH BRAND IN RUPEES IS SHOWN IN THE OUTPUT.

5) AVERAGE RATING PER BRAND

```
grunt> gbrand = GROUP project by brand;
2021-10-03 11:16:11,170 [main] WARN org.apache.pig.newplan.BaseOperatorPlan -
grunt> e = foreach gbrand generate group , ROUND(AVG(project.rating));
2021-10-03 11:16:24,271 [main] WARN org.apache.pig.newplan.BaseOperatorPlan -
grunt>
```

OUTPUT:

```
(Keya,4)
(Kiwi,4)
(LISO,1)
(Lays,4)
(MOOZ,4)
(Mars,5)
(Maxi,4)
(Maxo,4)
(Maya,)
(Moov,4)
(NOCD,4)
(Nice,5)
(Nike,4)
(Nova,3)
(Nyle,4)
(OGMO,3)
(Olay,4)
(Only,4)
(Ongs,4)
(Open,4)
(Opex,4)
(Oreo,4)
(PIPO,4)
(Plum,4)
(Polo,)
(Post,4)
(Pril,4)
(Pure,4)
(QIPS,5)
(QRAA,4)
(RYCA,)
(Rage,)
(Real,4)
(Red",)
(Rica,4)
(SOVI,5)
(Seni,4)
(Smac,4)
(Sofy,4)
(Solo,4)
(Tagz,4)
(Tang,4)
(Tirol,5)
```

AVERAGE RATING PER BRAND IS SEEN IN THE OUTPUT I.E OUTPUT RATINGS IN THE RANGE(1-5)

BRANDS LIKE OREO ,TANG HAVE AVERAGE RATING 4 AND BRANDS LIKE MARS AND NICE HAVE AVERAGE RATING 5

6) MOST PROFITABLE PRODUCT

```
grunt> h1 = group project all;
2021-10-03 12:04:47,111 [main] WARN org.apache.pig.newplan.B
grunt> I = foreach h1 generate MAX(project.Profit) as max;
2021-10-03 12:05:41,872 [main] WARN org.apache.pig.newplan.B
grunt> x = filter project by Profit==I.max;
2021-10-03 12:06:09,087 [main] WARN org.apache.pig.newplan.B
grunt> dump x;
```

OUTPUT:

```
2021-10-03 12:06:15,550 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat
2021-10-03 12:06:15,550 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil
("Premium Cloth Dryer/Drying Stand - Foldable, Silver, BB 1434", "Kitchen,,,DP,4649.0,8969.0,)
grunt> █
```

FROM THE OUTPUT WE CAN SEE THE MOST PROFITABLE PRODUCT IS PREMIUM CLOTH DRYER FROM THE BRAND DP

7) LEAST PROFITABLE PRODUCTS

```
grunt> h1 = group project all;
2021-10-03 12:00:15,950 [main] INFO org.apache.pig.newplan.B
g Thunderbird Mail ach h1 generate MIN(project.Profit) as min;
2021-10-03 12:24:49,976 [main] WARN org.apache.pig.newplan.B
grunt> x = filter project by Profit==I.min;
2021-10-03 12:27:27,121 [main] WARN org.apache.pig.newplan.B
grunt> dump x;
```

OUTPUT:

```
(Clothes Disinfectant & Refreshing Spray,Cleaning & Household,All Purpose Cleaners,Savlon,175.0,175.0,Disinfectant Spray & Cleaners,4.5,0.0,0.0)
(bb Home Floor Cleaner - Neem 500 ml + bb Home Floor Cleaner - Citrus 500 ml,Cleaning & Household,All Purpose Cleaners,bb Combo,196.0,196.0,Floor & Other Cleaners,4.1,0.0,0.0)
(Black Phenyl,Cleaning & Household,All Purpose Cleaners,Germt Chek,48.0,48.0,Floor & Other Cleaners,4.2,0.0,0.0)
(Perfumed Herbal Phenyle - White,Cleaning & Household,All Purpose Cleaners,Smart Home,60.0,60.0,Floor & Other Cleaners,4.1,0.0,0.0)
(Floor Cleaner - Pine White,Cleaning & Household,All Purpose Cleaners,Honey Home Fresh,50.0,50.0,Floor & Other Cleaners,4.0,0.0,0.0)
(Phenyle - Sensation (Herbal 2 in 1),Cleaning & Household,All Purpose Cleaners,Germisol,60.0,60.0,Floor & Other Cleaners,4.2,0.0,0.0)
(Platinum Active-Shield Disinfectant Toilet Cleaner Liquid - Marine,Cleaning & Household,All Purpose Cleaners,Harpic,99.0,99.0,Toilet Cleaners,4.3,0.0,0.0)
(Toilet Air Freshener - Lavender,Cleaning & Household,All Purpose Cleaners,Odonil ,49.0,49.0,Toilet Cleaners,4.1,0.0,0.0)
(Toilet Cleaner - Lemon,Cleaning & Household,All Purpose Cleaners,BB Home,168.0,168.0,Toilet Cleaners,4.1,0.0,0.0)
(Platinum Active-Shield Disinfectant Toilet Cleaner Liquid - Lavender,Cleaning & Household,All Purpose Cleaners,Harpic,99.0,99.0,Toilet Cleaners,4.3,0.0,0.0)
(Power Plus Disinfectant Toilet Cleaner,Cleaning & Household,All Purpose Cleaners,Harpic,356.0,356.0,Toilet Cleaners,4.4,0.0,0.0)
(Disinfectant Toilet Cleaner Pro - Colour Changing Formula - Kills 99.9% germs,Cleaning & Household,All Purpose Cleaners,BB Home,91.0,91.0,Toilet Cleaners,4.1,0.0,0.0)
(Dishwasher Salt,Cleaning & Household,All Purpose Cleaners,Fortune ,200.0,200.0,Imported Cleaners,5.0,0.0,0.0)
(Fabric Softener - Classic Blue,Cleaning & Household,All Purpose Cleaners,Comfort,450.0,450.0,Imported Cleaners,4.1,0.0,0.0)
(Bleach - Original,Cleaning & Household,All Purpose Cleaners,Clorox,550.0,550.0,Floor & Other Cleaners,4.4,0.0,0.0)
(Disinfecting Wipes - Fresh Scent,Cleaning & Household,All Purpose Cleaners,Clorox,450.0,450.0,Disinfectant Spray & Cleaners,4.6,0.0,0.0)
(Dishwasher Tablet,Cleaning & Household,All Purpose Cleaners,Fortune ,440.0,440.0,Imported Cleaners,5.0,0.0,0.0)
(Fabric Conditioner - Rose Fresh,Cleaning & Household,All Purpose Cleaners,Comfort,450.0,450.0,Imported Cleaners,3.0,0.0,0.0)
(Bleach - Lemon,Cleaning & Household,All Purpose Cleaners,Clorox,160.0,160.0,Floor & Other Cleaners,4.1,0.0,0.0)
(Floor Cleaner - Lime,Cleaning & Household,All Purpose Cleaners,Florona,70.0,70.0,Floor & Other Cleaners,3.8,0.0,0.0)
(Toilet Air Freshener - Rose,Cleaning & Household,All Purpose Cleaners,Odonil ,49.0,49.0,Toilet Cleaners,4.1,0.0,0.0)
(Disinfectant Bathroom Cleaner - Lemon,Cleaning & Household,All Purpose Cleaners,Harpic,44.0,44.0,Floor & Other Cleaners,4.1,0.0,0.0)
(Herbal Floor Cleaner,Cleaning & Household,All Purpose Cleaners,Nmyle,84.0,84.0,Floor & Other Cleaners,5.0,0.0,0.0)
(Rinse Aid,Cleaning & Household,All Purpose Cleaners,Fortune ,150.0,150.0,Imported Cleaners,3.0,0.0,0.0)
(Dishwasher Salt,Cleaning & Household,All Purpose Cleaners,Finish,340.0,340.0,Imported Cleaners,5.0,0.0,0.0)
(Disinfecting Wipes Fresh Scent,Cleaning & Household,All Purpose Cleaners,Clorox,250.0,250.0,Imported Cleaners,4.1,0.0,0.0)
(Bathroom Cleaner,Cleaning & Household,All Purpose Cleaners,BB Home,91.0,91.0,Toilet Cleaners,4.1,0.0,0.0)
(Disinfectant Bathroom Cleaner - Floral,Cleaning & Household,All Purpose Cleaners,Harpic,44.0,44.0,Floor & Other Cleaners,4.2,0.0,0.0)
(Visible Power Toilet Cleaner,Cleaning & Household,All Purpose Cleaners,Mr Muscle,75.0,75.0,Toilet Cleaners,4.2,0.0,0.0)
(Bleach - Lavender,Cleaning & Household,All Purpose Cleaners,Clorox,160.0,160.0,Floor & Other Cleaners,4.4,0.0,0.0)
(Proclean Floor Cleaner - Lavender,Cleaning & Household,All Purpose Cleaners,Godrej,179.0,179.0,Floor & Other Cleaners,4.5,0.0,0.0)
(Disinfecting Wipes Fresh Scent,Cleaning & Household,All Purpose Cleaners,Clorox,130.0,130.0,Imported Cleaners,4.1,0.0,0.0)
(Bleach - Lemon,Cleaning & Household,All Purpose Cleaners,Clorox,550.0,550.0,Floor & Other Cleaners,4.3,0.0,0.0)
(Herbal Floor Cleaner,Cleaning & Household,All Purpose Cleaners,Nmyle,156.0,156.0,Floor & Other Cleaners,4.1,0.0,0.0)
(Toilet Seat Sanitizer - Deodoriser,Cleaning & Household,All Purpose Cleaners,Sanitit,150.0,150.0,Toilet Cleaners,3.7,0.0,0.0)
(IMPORTED - 3Ply Kitchen Towel Super Absorbent Paper Tissue Roll,Cleaning & Household,All Purpose Cleaners,Selpak,1340.0,1340.0,Imported Cleaners,4.5,0.0,0.0)
(Heavy Starch - Fresh Lavender Scent,Cleaning & Household,All Purpose Cleaners,Faultless,570.0,570.0,Imported Cleaners,4.0,0.0,0.0)
(Clean-Up All-Purpose Cleaner with Bleach,Cleaning & Household,All Purpose Cleaners,Clorox,425.0,425.0,Floor & Other Cleaners,4.3,0.0,0.0)
(Gottle Tap Cleaner - Multipurpose Cleaner,Cleaning & Household,All Purpose Cleaners,GOTILE,375.0,375.0,Imported Cleaners,4.1,0.0,0.0)
(Ato Max Dishwasher Tablets - Regular,Cleaning & Household,All Purpose Cleaners,Finish,599.0,599.0,Imported Cleaners,4.1,0.0,0.0)
(Multipurpose Germ Protection Anti-Bacterial Wipes with Lemon Fragrance,Cleaning & Household,All Purpose Cleaners,Corsafe Advance,100.0,100.0,Disinfectant Spray & Cleaners,4.1,0.0,0.0)
(Dishwasher - Cleaner,Cleaning & Household,All Purpose Cleaners,Clean N Fresh,279.0,279.0,Floor & Other Cleaners,4.0,0.0,0.0)
(Dishwasher Tablets,Cleaning & Household,All Purpose Cleaners,Cleox,575.0,575.0,Imported Cleaners,1.0,0.0,0.0)
(Rinse Aid,Cleaning & Household,All Purpose Cleaners,Clean N Fresh,249.0,249.0,Imported Cleaners,4.1,0.0,0.0)
(Dishwasher Tablets,Cleaning & Household,All Purpose Cleaners,Cleox,1400.0,1400.0,Imported Cleaners,4.1,0.0,0.0)
(Polish Warrish Remover - Silver Wipes,Cleaning & Household,All Purpose Cleaners,Meitner,650.0,650.0,Imported Cleaners,4.1,0.0,0.0)
(Cleaning Acid,Cleaning & Household,All Purpose Cleaners,Zermitol,58.0,58.0,Floor & Other Cleaners,4.2,0.0,0.0)
(Floor Cleaner - Sandal,Cleaning & Household,All Purpose Cleaners,Florona,300.0,300.0,Floor & Other Cleaners,3.9,0.0,0.0)
```

FROM THE OUTPUT WE CAN SEE THE PRODUCTS WITH 0 RUPEES PROFIT AND THESE ARE MOSTLY THE PRODUCTS THAT FALL IN THE CATEGORY OF HOUSEHOLD/CLEANING ETC.

8) MOST EXPENSIVE PRODUCT

```
grunt> h1 = group project all;
2021-10-03 12:39:23,643 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning
IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> I = foreach h1 generate MAX(project.sale_price) as exp;
2021-10-03 12:39:50,249 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning
IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> x = filter project by sale_price==I.exp;
2021-10-03 12:40:09,023 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning
IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> dump x;
```

OUTPUT:

```
- Total input paths to process : 1
(Epilator SE9-9961 Legs-Body-Face,Beauty & Hygiene,Feminine Hygiene,Braun,8184.44,10769.0,Hair Rem
oval,4.1,2584.56,31.6)
grunt>
```

FROM THE OUTPUT WE CAN SEE EPILATOR IS THE MOST EXPENSIVE PRODUCT AS PER SALE PRICE.

9) CHEAPEST PRODUCT

```
kevin@kevin-VirtualBox: ~/Desktop
2021-10-03 12:48:00,034 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> I = foreach h1 generate MIN(project.sale_price) as cheap;
2021-10-03 12:48:17,939 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> x = filter project by sale_price==I.cheap;
2021-10-03 12:48:25,078 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> dump x;
```

OUTPUT:

```
cc - Total input paths to process : 1
(Curry Leaves,Fruits & Vegetables,Herbs & Seasonings,Fresho,2.45,15.0,Indian & Exotic Herbs,4.1,12.55,512.2)
grunt>
```

FROM THE OUTPUT THE CHEAPEST PRODUCT IS FOUND TO BE CURRY LEAVES

10)NUMBER OF PRODUCTS PER TYPE

```
kevin@kevin-VirtualBox: ~/Desktop
2021-10-03 13:07:06,132 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> rs = foreach tg generate group , COUNT(project);
2021-10-03 13:07:32,149 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_DOUBLE 4 time(s).
grunt> dump rs;
```

OUTPUT:

(Antiseptics & Bandages,25)
(Baby Dishes & Utensils,7)
(Cooking & Baking Needs,5)
(Cream & Cheese Spreads,14)
(Electronic Accessories,12)
(Exam Pads & Pencil Box,21)
(Exotic & Flavoured Tea,47)
(Extra Virgin Olive Oil,50)
(Flavoured & Other Oils,51)
(Floor & Other Cleaners,113)
(Floss & Tongue Cleaner,3)
(Fresh Salads & Sprouts,10)
(Gourmet Tea & Tea Bags,295)
(Hair & Scalp Treatment,67)
(Hand Wash & Sanitizers,206)
(Heat & Eat Ready Meals,134)
(Moustache & Beard Care,69)
(Other Baby Accessories,3)
(Pet Food & Accessories,3)
(Shower Gel & Body Wash,125)
(Sports & Energy Drinks,19)
(Supplements & Proteins,207)
(Breakfast & Snack Mixes,201)
(Cream Biscuits & Wafers,94)
(Dishwash Bars & Powders,29)
(Glucose & Milk Biscuits,18)
(Gourmet Juices & Drinks,92)
(Maternity Personal Care,15)
(Regular & White Vinegar,5)
(Soap Cases & Dispensers,10)
(Appliances & Electricals,13)
(Balsamic & Cider Vinegar,35)
(Body Scrubs & Exfoliants,10)
(Disposable Cups & Plates,16)
(Flavoured & Greek Yogurt,22)
(Imported Oats & Porridge,51)
(International Chocolates,103)
(Namkeen & Savoury Snacks,265)
(Organic & Cold Press Oil,16)
(Tampons & Menstrual Cups,26)
(Dhatu Organics & Naturals,3)
(Dishwash Liquids & Pastes,39)
(Dry Shampoo & Conditioner,8)
(Face Masks & Safety Gears,34)
(Maternity Health Supplements,5)
(Disinfectant Spray & Cleaners,62)

FROM THE OUTPUT WE CAN SEE COUNT BY TYPE

FACE CARE , SHAMPOO ,HANDWASH HAVE HIGH NUMBERS.

CONCLUSION:

MOST PROFITABLE CATEGORIES ARE:

BEAUTY & HYGIENE

Comparatively fruits and vegetables and cleaning products are least profitable

So marketing and advertising or discounts can be added

And low rated products like cleaning and household should be also taken care using similar strategies in order to increase sales and profit.